

# Rules for evaluating the difficulty of memory problems

CATHERINE A. HALE and ROBERT KAIL  
Purdue University, West Lafayette, Indiana

Subjects were shown pairs of photographs; each photograph depicted an individual trying to remember the names of animals. The individual depicted was either an adult or a child, and the number of animals to be remembered varied from 1 to 20. The subjects judged which individual in each pair of photographs would be less likely to recall all the animals. Judgments for most children in Grades 1 through 6 conformed to a rule in which the principal determinant of difficulty was the number of animals. In contrast, most eighth-graders and college students used a rule in which the separate and interactive effects of the number of animals and the age of the depicted individual were considered.

As children develop, they are more likely to use strategies to aid retention. One suggested explanation for this developmental change is that older children are better able to gauge the difficulty of a memory task and hence are more likely to select a task-appropriate mnemonic. In fact, there is much research indicating that older children better understand individual variables that influence the difficulty of memory tasks (Flavell & Wellman, 1977). For example, 4-year-olds know that the number of items to be remembered contributes to task difficulty, but not until age 8 or 9 years do children understand that the presence of semantic relations among items to be remembered can simplify a memory task (Flavell & Wellman, 1977).

The purpose of the present study was to identify the specific rules that individuals use to determine task difficulty when more than a single memory variable is relevant. The task presented to the subjects in the present experiment involved pairs of photographs; each photograph depicted an individual looking intently at pictures of animals. The photographs in each pair were varied systematically along two dimensions: (1) the age of the individual, who was either a 9-year-old or an adult; and (2) the number of animals, which ranged from 1 to 20. In all other respects, the photographs in each pair were identical. The subjects were shown a pair of photographs and were asked to select the photograph in which it would be less likely that the depicted individual would remember all of the animals. Age and number were selected as the two memory-relevant variables, because,

We would like to thank: Marilyn Nippold for her help in testing children; administrators, teachers, and students at the Dayton and St. Lawrence Elementary Schools for their friendly cooperation throughout this research; and Stu Offenbach and Diane Collier for their comments on a previous draft of this manuscript. Catherine Hale is now at St. Mary's College, South Bend, Indiana. Requests for reprints should be sent to Robert Kail, Department of Psychological Sciences, Purdue University, West Lafayette, Indiana 47907.

by age 5 years, children know the effects of these variables in isolation (e.g., Flavell & Wellman, 1977); hence, it is plausible that young children would attempt to consider both of these variables.

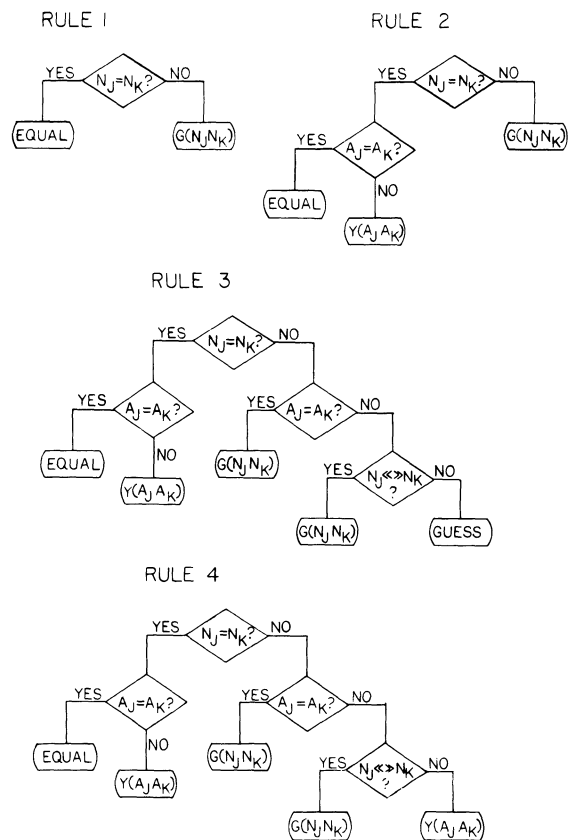


Figure 1. Four rules for judging the more difficult of two memory problems.  $N_J$  and  $N_K$  refer to the number of items in the two problems;  $A_J$  and  $A_K$  refer to the ages of the individuals in the problems;  $G(N_J, N_K)$  means "the greater of  $N_J$  and  $N_K$ " and  $Y(A_J, A_K)$  means "the younger of  $A_J$  and  $A_K$ ."

Table 1  
Types of Problems and Predicted Responses for Rules 1-4

Type of Problem	Four Instances of the Problem	Predicted Accuracy for Rule (in Percent)			
		1	2	3	4
Age	9A/9C, 10A/10C, 11A/11C, 12A/12C	0	100	100	100
Number	8C/10C, 8A/10A, 7C/9C, 7A/9A	100	100	100	100
Conflict-Number	1C/17A, 2C/18A, 3C/19A, 4C/20A	100	100	100	100
Conflict-Age	8C/9A, 9C/10A, 10C/11A, 11C/12A	0	0	50	100

Note—In the instances, the numeral refers to the number of animals in the set, whereas the letter refers to the age of the individual (A, adult; C, child). Also, the instances have been presented in a way that highlights the similarities of instances within each problem type. The letters and numbers to the left and right of the diagonal do not correspond to the left/right presentation of the pictures on the cards.

Rational analyses, preliminary work with children and adults, and existing research on scientific-reasoning tasks (Siegler, 1981) suggested four potential rules for solving these problems (Figure 1). To differentiate among these rules, four types of problems were created (Table 1). In number problems, the individuals in the two photographs were the same age (i.e., both children or both adults), but the photographs differed in the number of items to be remembered. In age problems, the number of items was constant, whereas age varied. In conflict problems, both age and number of items differed. In conflict-number problems, the numbers of animals in the two sets always differed by 16; in conflict-age problems, the numbers of animals differed by 1.

The four rules depicted in Figure 1 lead to different judgments on the problems presented in Table 1. Hence, if younger children are more likely to use simpler rules, then distinct age-related patterns of performance are predicted for the four types of problems: (1) performance on number and conflict-number problems should not vary with age, for all rules lead to correct judgments on these problems; (2) performance on age problems should increase moderately with age, because Rules 2-4 lead to a correct judgment, whereas Rule 1 results in incorrect judgments; and (3) accuracy on conflict-age problems should increase markedly with age, because Rules 1 and 2 lead to wrong judgments, whereas Rules 3 and 4 lead to correct responses.<sup>1</sup>

## METHOD

### Participants

The participants were 10 first-graders [median age: 6 years; 6 months (6;6)], 25 second-graders (7;8), 18 third-graders (8;9), 31 fourth-graders (9;11), 12 fifth-graders (10;8), 26 sixth-graders (11;11), 13 eighth-graders (13;9), and 20 college undergraduates (19;11). Approximately 62% of the participants were males.

### Memory Problems

The stimuli for the memory judgments were 16 pairs of 9 x 13 cm color photographs; each photograph was mounted on a piece of black cardboard. Each photograph depicted a person looking at pictures of animals, as if studying them. The photographs differed with respect to the age of the individual, who was either a 9-year-old boy or girl or a 30-year-old man or woman, and to the number of animals, which ranged from 1 to 20. The individual's face and trunk were quite clear in the

photographs, as was the number of different pictures, but the individual animals were not.

The 16 pairs of memory problems listed in Table 1 were selected from a larger pool of 24 problems that had been administered, in written form, to eight developmental psychologists. Four examples of each of the four types of problems listed in Table 1 were selected; the criterion used was that all eight judges had chosen one of the two problems in a pair as being the more difficult. Photographs corresponding to the 16 pairs of memory problems were then prepared. The more difficult problem appeared equally often on the left and right sides of the cards.

### Procedure

Each participant was tested individually, in a session lasting approximately 15 min. First, the experimenter explained the nature of the judgments that were to be made. The participants were shown, in succession, pairs of 9 x 13 cm color photographs, each photograph depicting a young girl about to jump toward a ball placed on a lawn. The photographs in each pair were identical except that (1) in one pair, the ball was approximately 1 m from the girl in one photograph and 7 m away in the second photograph, and (2) in the second pair, the girl was approximately 1 m from a volleyball in one photograph and 1 m from a football in the second. The participants were asked to judge which girl would have greater difficulty in jumping over the ball. All participants responded correctly on both problems.

A recall task followed in which three cards—depicting a tiger, a monkey, and an elephant—were placed in a row in front of the participant, who was asked to remember the animals. After 5 sec, the cards were removed and the participant was asked to recall the animals. Every subject recalled all three animals. Finally, the 16 memory problems were presented twice, for a total of 32 problems. The experimenter explained that the people depicted in the photographs had to remember the items shown, exactly as the participant had done in the recall task just completed. For each card, the participant was asked which of the two persons in the photographs would find it more difficult to remember all the animals or whether the two persons would do about the same. All 16 problems were presented individually in this manner, in a randomly determined order. After the 16 problems had been presented and the subject's judgments recorded, the cards were shuffled and presented a second time.

## RESULTS

### Judgments

Correct responses increased with grade [ $F(7,147) = 4.25, p < .001$ ] and differed across types of problems [ $F(3,441) = 375.83, p < .001$ ]. Of greater interest was the significant interaction between grade and type of problem [ $F(21,441) = 7.1, p < .001$ ], which reflected

Table 2  
Percentage of Individuals at Four Grade Levels Using Different Rules

Grade Level	N	Rule						Unclassified
		1	1'	2	2'	3	4	
1 and 2	35	2.9	5.7	34.3	5.7	5.7	2.9	42.9
3 and 4	49	4.1	8.2	32.7	0	8.2	6.1	40.8
5 and 6	38	15.8	5.3	23.7	2.6	10.5	5.3	36.8
8 and college	33	0	3.0	12.1	0	0	54.5	30.3

Note—Due to rounding, the percentages in each row may not sum to 100%.

the pattern of results predicted on the assumption that the use of Rules 1-4 is grade related. Accuracy on number and conflict-number problems did not vary with grade level ( $F_s \leq 1$ ), whereas performance on age and conflict-age problems increased with grade ( $F_s \geq 5.05$ ,  $p < .001$ ). In addition, accuracy on conflict-age problems was significantly below chance level ( $p < .05$ ) for Grades 1-6, not different from chance at Grade 8, and significantly above chance ( $p < .05$ ) only for undergraduates.

Use of Rules

For a person to be classified as having used a particular rule (see Table 2), 24 of 32 responses had to conform to those predicted by the rule, as did 6 of 8 responses on each type of problem. With these criteria, the likelihood that a person's responses would conform by chance alone to a pattern predicted by a rule is less than  $2 \times 10^{-6}$ . Approximately 62% of the subjects responded in patterns corresponding to one of the rules, and the proportion of subjects using rules did not vary significantly with grade level [ $\chi^2(3) = 1.45$ ].

Two new rules that appear in Table 2—1' and 2'—are unanticipated variants of Rules 1 and 2. The decision phase of Rule 1' corresponds exactly to Rule 1, as does the response rule when the numbers of items in the two memory problems are not equal. However, when the items are equal, as in age problems, subjects using Rule 1' guess, resulting in 50% accuracy on these problems. On the remaining three types of problems, their responses are identical to those of subjects using Rule 1. Rule 2' differs from Rule 2 in a similar manner. All decision processes shown in Figure 1 are alike for Rules 2 and 2'. The sole difference between the rules is that, when the numbers of items are equal but ages are not (i.e., age problems), subjects using Rule 2' predict that adults will have greater difficulty. Hence, persons using Rule 2' err on all age problems, but otherwise respond like individuals using Rule 2.

Of particular interest are developmental changes in the use of rules. Rules 2 and 4 were the most commonly used rules. However, use of Rule 2 declined at the higher grade levels, whereas use of Rule 4 increased [ $\chi^2(3) = 31.23$ ,  $p < .001$ ]. Use of Rule 2 or 4 and grade level were strongly associated (Cramer's  $\phi = .68$ ). Comparisons between adjacent pairs of grades (with Yates's corrected chi-squares) revealed only one significant

difference, between fifth- and sixth-graders on the one hand, and eighth-graders and undergraduates on the other [ $\chi^2(1) = 11.05$ ,  $p < .001$ ].<sup>2</sup>

DISCUSSION

The present research was designed to reveal rules that individuals use to determine the difficulty of memory problems. The principal results were that: (1) a majority of the individuals at each grade used one of the proposed rules; (2) each of the rules was used by some individuals; and (3) Rule 2 was the modal rule for children, and Rule 4 was the modal rule for adolescents and young adults.

There are some important limiting conditions on the rule system and the supporting data presented here. First, we have not specified the precise manner in which individuals judge the equality of items in two sets (or the ages of two individuals) and have not determined what would constitute a "substantially greater" difference (as in Rule 4). Specifying these processes is a goal for future research and would, presumably, result in replacing many of the decision phases in Rules 1-4 with a series of more specific operations. Second, the problems formulated to test Rules 1-4 represent a limited set of ages and items to be remembered. Using a wider range of problems would test the generality of the rules, and also would be useful in determining the specific processes underlying judgments of age and memory-relevant numerosity.

Finally, as noted in the introduction, Rules 1-4 were derived in part from work on the development of scientific reasoning. Generic versions of the four rules depicted in Figure 1 characterize performance on numerous reasoning tasks. Specifically, developmental change on many reasoning and problem-solving tasks consists of the acquisition of rules of successively greater scope, at rates that vary from task to task (Siegler, 1981). Our findings provide additional evidence for the generality of this developmental sequence in the use of rules, by showing that the rules apply to children's evaluation of memory problems as well as to scientific reasoning tasks.

REFERENCES

FLAVELL, J. H., & WELLMAN, H. M. (1977). Metamemory. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition*. Hillsdale, NJ: Erlbaum.  
SIEGLER, R. S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development*, 46(2, Serial No. 189).

NOTES

1. In these four rules, judgments of number take precedence over judgments of age. One could just as easily postulate a complementary set of four rules in which age takes precedence: This would amount to replacing all judgments about number in Figure 1 with judgments about age, and vice versa. Thus, in this version of Rule 1, persons would judge the difficulty of memory

problems solely in terms of the memorizer's age; the number of items would never be considered. Neither the children in our preliminary testing nor the participants in the present study mentioned such rules or responded in a way that was consistent with them. This complementary set of rules does explain, however, the presence of number and conflict-number problems. Although these problems do not distinguish the four rules depicted in Figure 1, they are needed to distinguish the complementary four "age-precedence" rules: With this complementary set of rules, age and conflict-age problems are always solved correctly, whereas number and conflict-number problems are not.

2. A sizable minority of subjects could not be classified. One possibility is that these individuals used rules not discernible to

us. A second possibility is that they used Rules 1-4 less skillfully than did their peers, with the result that their responses were too inconsistent to meet our criteria for classifying subjects. Support for the second position comes from the fact that, of the 59 unclassified individuals, 25 missed adherence to Rule 2 by a single response (2, 4, 2, 7, 3, 4, 2, and 1, from Grades 1-6, Grade 8, and college, respectively) and 8 missed Rule 4 by one response (1 each from Grades 2 and 4; 2 each from Grades 3 and 8 and college). If these individuals are classified as rule users, then the system of rules presented here accounts for the responses of approximately 83% of the participants.

(Manuscript received for publication November 8, 1983.)